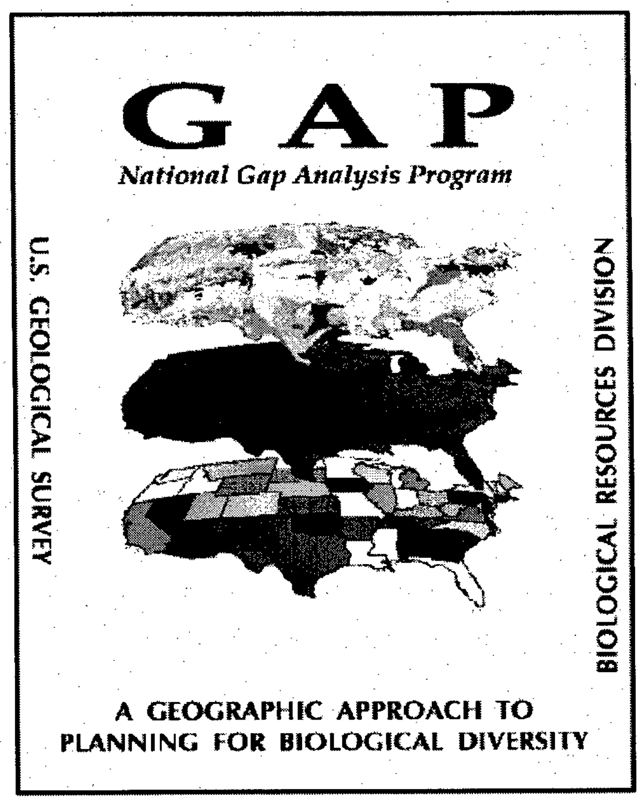
USGS
science for a changing world

98-G001

Upper Midwest Gap Analysis Program Image Processing Protocol



U.S. Department of the Interior
U.S. Geological Survey
Environmental Management Technical Center

June 1998

The Environmental Management Technical Center was established in 1986 as a center for ecological monitoring and analysis of the Upper Mississippi River System.

U.S. Geological Survey
Environmental Management
Technical Center

CENTER DIRECTOR
Robert L. Delaney

GEOSPATIAL APPLICATIONS
ACTING DIRECTOR
Norman W. Hildrum

PROGRAM OPERATIONS
ACTING DIRECTOR
Linda E. Leake

REPORT EDITOR
Jerry Cox

Printed on recyled paper

Upper Midwest Gap Analysis Program
Image Processing Protocol

by

Thomas Lillesand (Protocol Project Coordinator) and Jonathan Chipman (Protocol Editor)
Environmental Remote Sensing Center
University of Wisconsin–Madison
1225 West Dayton Street
Madison, Wisconsin 53706-1695

David Nagel, Heather Reese, Matthew Bobo, and Robert Goldmann (Major Contributors)
Wisconsin Department of Natural Resources
PO Box 7921
Madison, Wisconsin 53707

June 1998

U.S. Geological Survey
Environmental Management Technical Center
575 Lester Avenue
Onalaska, Wisconsin 54650

Suggested citation:

Lillesand, T., J. Chipman, D. Nagel, H. Reese, M. Bobo, and R. Goldmann. 1998. Upper Midwest Gap analysis program image processing protocol. Report prepared for the U.S. Geological Survey, Environmental Management Technical Center, Onalaska, Wisconsin, June 1998. EMTC 98-G001. 25 pp. + Appendixes A–C

Additional copies of this report may be obtained from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161 (1-800-553-6847 or 703-487-4650). Also available to registered users from the Defense Technical Information Center, Attn: Help Desk, 8725 Kingman Road, Suite 0944, Fort Belvoir, VA 22060-6218 (1-800-225-3842 or 703-767-9050).

United States Department of the Interior
U.S. Geological Survey
Environmental Management Technical Center
575 Lester Avenue, Onalaska, Wisconsin 54650-8552

IN REPLY REFER TO:

July 15, 1998

Authors
Reviewers
Science Review Committee
U.S. Geological Survey Science Centers and Regional Offices

Dear Colleague:

I am enclosing the U.S. Geological Survey publication "Upper Midwest Gap Analysis Program Image Processing Protocol" by Thomas Lillesand and Jonathan Chipman of the Environmental Remote Sensing Center, University of Wisconsin–Madison, and David Nagel, Heather Reese, Matthew Bobo, and Robert Goldmann of the Wisconsin Department of Natural Resources. The Environmental Management Technical Center provides coordination for the Upper Midwest Gap Analysis Program.

Please contact me at (608) 783-7550, extension 51, if you have questions regarding this report.

Sincerely,

Robert L. Delaney
Center Director

Enclosure:
98-G001

Contents

Figures

Preface

The Gap Analysis Program (GAP) is a U.S. Geological Survey project being implemented nationwide with the help of more than 400 cooperators, including the private sector, nonprofit organizations, and government agencies. The purpose of GAP is to identify gaps in the network of conservation lands with respect to land cover or habitat types as well as individual vertebrate species and to build partnerships around the development and application of this information (Scott et al. 1993).

Gap Analysis is conducted by combining the distribution of actual natural vegetation, mapped from satellite imagery and other data sources, with distributions of vertebrate and other taxa as indicators of biodiversity. The data are manipulated and displayed using computerized geographic information systems. Maps of species-rich areas, individual species of concern, and overall vegetation types are generated. Using geographic information systems, this information can be analyzed to show where land-based conservation efforts need to be focused to achieve conservation of overall biodiversity most efficiently.

The U.S. Geological Survey Environmental Management Technical Center facilitates the Upper Midwest GAP (UMGAP), a cooperative effort with the states of Illinois, Michigan, Minnesota, and Wisconsin. Mapping support is also provided to the states of Indiana and Iowa in an effort to produce a common database for the Upper Midwest region.

The protocol describes both the underlying philosophy and the operational details of the land cover classification activities being performed as part of UMGAP. Topics discussed include the hierarchical classification scheme, ground reference data acquisition, image stratification, and classification techniques. This discussion is primarily aimed at the image processing analysts involved in the UMGAP land cover mapping activities as well as others involved in similar projects. It is a "how-to" technical guide of interest to people responsible for satellite image processing.

Upper Midwest Gap Analysis Program Image Processing Protocol

by

Thomas Lillesand, Jonathan Chipman, David Nagel,
Heather Reese, Matthew Bobo, and Robert Goldmann

Abstract

This document presents a series of technical guidelines by which land cover information is being extracted from Landsat Thematic Mapper data as part of the Upper Midwest Gap Analysis Program (UMGAP). The UMGAP represents a regionally coordinated implementation of the national Gap Analysis Program in the states of Michigan, Minnesota, and Wisconsin; the program is led by the U.S. Geological Survey, Environmental Management Technical Center.

The protocol describes both the underlying philosophy and the operational details of the land cover classification activities being performed as part of UMGAP. Topics discussed include the hierarchical classification scheme, ground reference data acquisition, image stratification, and classification techniques. This discussion is primarily aimed at the image processing analysts involved in the UMGAP land cover mapping activities as well as others involved in similar projects. It is a "how-to" technical guide for a relatively narrow audience, namely those individuals responsible for the image processing aspects of UMGAP.

1. Introduction

Studies at the University of Wisconsin–Madison Environmental Remote Sensing Center and the Wisconsin Department of Natural Resources have led to the development of a proposed methodology for large-area land cover classification using satellite imagery. This protocol is intended to guide image processing analysts working on the combined statewide land cover mapping efforts of the Wisconsin Initiative for Statewide Cooperation on Landscape Analysis and Data (WISCLAND) and the Wisconsin portion of the Upper Midwest Gap Analysis Program (UMGAP). The Upper Midwest Gap Analysis Program represents a regionally coordinated implementation of the national Gap Analysis Program (GAP) in the states of Michigan, Minnesota, and Wisconsin, led by the U.S. Geological Survey (USGS), Environmental Management Technical Center. The image processing procedures developed for WISCLAND, developed specifically for Wisconsin, form the general basis for the UMGAP image processing activities being applied simultaneously in Michigan and Minnesota. The latter two states, however, are making appropriate modifications to the protocol to reflect local programmatic interests and preexisting geographic information systems data sources.

The protocol describes the underlying philosophy and operational details of the land cover classification activities being performed as part of UMGAP. The hierarchical classification scheme is described first, followed by the ground reference data collection process. A stratified sampling scheme is used to acquire ground reference data for training purposes. Prior to classification, Landsat Thematic Mapper (TM) satellite images are stratified according to several factors, and individual strata are classified separately. The primary classification method used here is "guided clustering," a hybrid technique combining elements of both supervised and unsupervised classification methods. The overall genesis of these classification guidelines can be found in Lillesand (1994).

This discussion is aimed at a relatively narrow audience, that is the image analysts responsible for actually performing the image classification involved in the above land cover mapping programs as well as others involved in similar projects. Accordingly, this document focuses on the "how-to" technical steps necessary

to effect the image processing (and related geographic information systems analyses) being employed in UMGAP; for this reason, portions of this document include references to specific ERDAS Imagine and ARC/INFO commands and processes.[1] Also, the methods described herein are the result of ongoing studies, and many of these procedures are evolving as they are exercised in a production environment.

2. Selection of an Extendable Coding Scheme

One of the most important and difficult steps in planning a land cover classification project is selection of the categories to be discriminated in the mapping effort. The classification scheme should be compatible with existing national systems and yet represent local land cover characteristics. Selecting the appropriate level of categorical detail is also important. Choosing an overabundance of categories can lead to considerable confusion among cover types, whereas selecting too few classes may not meet user needs.

With these considerations in mind, a considerable effort was made to develop a classification scheme that was (1) compatible with existing national schemes, (2) reflective of Upper Midwest cover types, (3) realistic in terms of the TM sensor capabilities, considering that some ancillary data would also be used to aid the classification process, and (4) extendable under ideal classification conditions or with an improvement in technology. To accomplish this task, a classification scheme committee of WISCLAND participants was formed representing the Wisconsin Department of Natural Resources, the Environmental Remote Sensing Center, the U.S. Forest Service, and the USGS.

Numerous existing classification schemes were studied to help guide the structure and categorical detail of the UMGAP scheme. Some of these include "A Land Use and Land Cover Classification System for Use With Remote Sensor Data" (Anderson et al. 1976), "A Modified Wetland/Upland Land Cover Classification System for Use With Remote Sensor Data" (Klemas et al. 1992), "A Coastal Land Cover Classification System for the NOAA Coastwatch Change Analysis Project" (Klemas et al. 1993), and "Midwest Regional Community Classification" (Faber-Langendoen 1993).

To develop a classification scheme representative of Upper Midwest cover types and reflective of TM sensor capabilities, a collection of works comprising published research and graduate theses was examined. Results from 12 studies, consisting of 31 separate classifications conducted in the Great Lakes region, were compiled into a single document. Accuracy figures for each land cover class in conjunction with category specificity were noted for each study. From these observations, a group of base categories was identified for inclusion in the UMGAP classification scheme, and additional extended categories were noted for possible use under ideal classification conditions, with improved technology, or through the inclusion of other data sources. These base and extended categories are listed in Appendix A, and definitions are included in Appendix B.

The national GAP standards (Jennings 1994) involve classification to the alliance level and consistency with the United Nations Educational, Scientific, and Cultural Organization/The Nature Conservancy system (United Nations Educational, Scientific, and Cultural Organization 1973), with certain limitations. Many of the UMGAP categories listed in Appendix A can be matched directly to individual alliances. Some categories, however, represent components of multiple alliances. For example, the classification system in Appendix A lists separate categories for beech, sugar maple, red maple, and three oak species; these represent several alliances including "beech-sugar maple" and "beech-oak-maple." At the 30- × 30-m

[1]References to these commands and processes are provided to clarify certain aspects of the protocol, and mention of particular software packages is not intended to express or imply the endorsement of same.

(0.09 ha) spatial resolution required by many end users of the UMGAP land cover data, the individual categories listed in Appendix A will be used. During the aggregation from the 0.09 ha initial classification to the final 100-ha GAP minimum mapping unit, the categories will be modified to reflect the standard GAP classes (see Section 6, Post-Classification Processing).

2.1 The Upper Midwest Gap Analysis Program Classification System

The classification system is hierarchical in character (i.e., more detailed classes can be collapsed into more general ones). For example, the extended class of "Orchard" can be generalized up one level to "Woody" or two levels to "Agriculture." The classification system is designed with an eye towards "crosswalking" it to other systems where possible. Whereas the system fully exploits the potential of automated image classification, it also recognizes its limitations. It is envisioned that the system can and will be extended through the use of additional land cover categories and other information sources. It provides a point of departure for such applications as GAP analysis. The need for potential extension, however, was recognized from the outset.

3. Ground Reference Data

Ground reference or groundtruth data must be collected to train the computer to recognize the various land cover categories latent in the TM imagery and to assess the categorical accuracy of the resulting classification. Ground reference data generally cannot be collected for large portions of the entire project area; therefore, representative samples are frequently used (Lillesand and Kiefer 1994). Several criteria must be considered when evaluating the suitability of any ground reference data set for land cover classification. First, the data collection method should be systematic, that is, representative of the entire area to be classified. Second, the method must have an element of randomness to avoid selection bias (Ott 1988). Third, a sufficient number of reference samples must be utilized to provide an appropriate sample density and ensure that the classification accuracy is known within a specified confidence level (Thomas and Allcock 1984). Fourth, the reference data must be reasonably contemporary with respect to the acquisition date of the imagery. Fifth, the level of accuracy of the reference data must be high. Last, the classification scheme used for collection of ground reference data must be compatible with the intended image processing classification system.

The UMGAP project includes both the collection of new ground reference data and the incorporation of preexisting reference data sets. For some areas of the region, particularly public lands, adequate ground reference data sets already exist that may meet the requirements for use in training and accuracy assessment. Also, for agricultural areas, previously collected data from the same year as the satellite imagery will be used. For other areas, new reference data will be collected in the field. The collection of new data in the field is described in Section 3.2, Nonagricultural Sample Site Selection and Training. The use of preexisting data is described in Section 3.3, Agricultural Sample Site Selection and Training.

To meet the six criteria outlined above, studies were conducted at the Wisconsin Department of Natural Resources and the Environmental Remote Sensing Center to examine methods for collecting and incorporating ground reference data. These studies were aimed at developing a sampling methodology whereby training and accuracy assessment data are collected simultaneously. Among the advantages of this strategy are the following: (1) redundant field work and data handling are minimized, (2) no changes occur on the ground between acquisition of training data and accuracy assessment data, and (3) discrepancies in the application of the classification system are avoided.

3.1 Sampling

3.1.1 Choosing Appropriate Ground Coverage

The first step in developing a sampling scheme was to determine the amount of ground area that should be sampled to include an adequate number of polygons for each land cover category. A statewide, completely randomized sampling scheme would require field staff to cover more ground than necessary to accurately represent all land cover categories. Because aerial photography is readily available for the region, and State Department of Natural Resources and other field staff cooperators are skilled in using this medium for navigation and interpretation, it was decided that aerial photos would serve as a base for delineating polygons for ground verification. The extent of individual photos would serve as a logical unit for sampling, thus restricting the ground area covered by field staff.

However, the data collection methods described here involve tradeoffs. These methods should produce a set of reference data representative of the full range of spectral variability present in each satellite image, thus providing ample training data for classification. On the other hand, the nonrandom aspects of the sampling scheme affect the use of these data for certain accuracy assessment purposes. This is discussed in Section 7.2, Thematic Accuracy Considerations.

Two large-area studies in the Great Lakes region by Luman (1992) and Bauer et al. (1994) were examined to help determine the number of photos that should be sampled to adequately represent all cover types. In addition, a pilot project examined previously classified TM scenes centered on various locations throughout Wisconsin. These data were processed by graduate students for various research projects conducted at the Environmental Remote Sensing Center. Four TM classifications capturing agricultural and forested regions of the state were subset in 2,048 x 2,048 pixel arrays and overlaid with a grid representative of 1:20,000 scale photo boundaries. Each photo covered about 4.5 km on a side. The 2,048 x 2,048 pixel array represented approximately 3,775 km^2, the size of a typical county in Wisconsin. The 1:20,000 scale photography was chosen because it was widely available and could be used as a surrogate for another readily available photo source, 1:40,000 scale National Aerial Photography Program (NAPP) frames.

Examination of the photography grid overlaid on the classified imagery suggested that a sample of about 6% of the photographs would capture enough variability in the scene to represent all but the least frequently present classes. To account for these rare categories, a sample of approximately 50% of the photography frames would be needed, which would involve a cost disproportionate to the importance of the infrequent categories. Other methods will be required to improve the representation of these infrequent categories.

Because current 1:40,000 scale NAPP photography is available to all three states involved in the UMGAP initiative, this product was used rather than the 1:20,000 scale photography. The 6% coverage deemed necessary could easily be transferred to the NAPP frames because a 1:20,000 scale photo covers one quarter of the area of a NAPP photo. The NAPP also has an advantage in that frames are centered on each of the four quarters of the 1:24,000 scale (7.5 min) USGS quadrangle maps ("quarter quads," Figure 1). This allows easy georeferencing of the photo frames in a geographic information system (GIS). In addition, because the NAPP photos cover four times the area of the 1:20,000 scale photos, more opportunities are offered to sample infrequently occurring categories.

Using NAPP photography, the fundamental sampling unit consists of one quarter of a photo, also referred to here as a USGS quarter quarter quadrangle (QQQ). Implementation of the sampling scheme is described below.

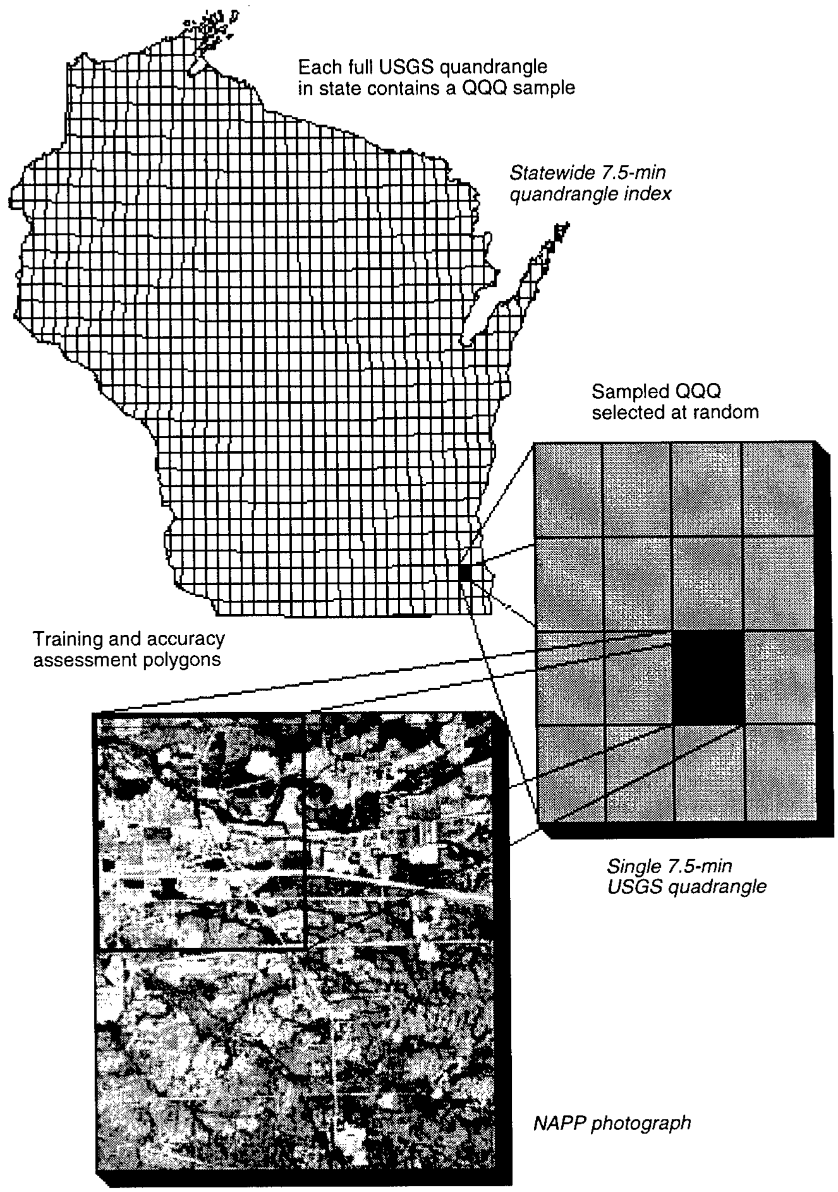


Figure 1. Geographically stratified sampling scheme.

3.1.2 Quarter Quarter Quadrangle Sampling Scheme

Completely randomized designs provide the ideal statistical basis for accuracy assessment but can prove impractical to implement (Congalton 1991), whereas a systematic approach is easier to implement but might not be acceptable for accuracy assessment (Congalton 1988). Thus, Congalton (1991) suggests that a combination of the random and systematic approaches be used for selecting samples. For the UMGAP project, a stratified scheme with random eastings and northings was chosen for selecting QQQs in which to delineate ground reference samples. The design allows for an essentially even distribution of sampling units throughout the state. A random north-south and east-west position is applied to each row and column of quad

sheets to minimize the effect of periodicity in the landscape. Berry and Baker (1968) suggest that this type of scheme is preferred for most land cover investigations, especially when underlying serial correlations (spatial autocorrelation) are unknown.

The sampling scheme is implemented as follows. Each USGS quad in the state, representing a primary cell or sampling stratum, is divided into four columns and four rows resulting in 16 secondary cells, each representing a QQQ. At random, a number (1–4) is assigned to each column and each row of primary cells. The random column assignment represents the north-south position for the secondary cell to be selected and the row assignment represents the east-west secondary cell position. A QQQ then is selected for each quadrangle based on the north-south and east-west random numbers generated (Figure 2).

For example, the northwest primary cell in Figure 2 has a north-south random number of 1 and an east-west assignment of 2. These random selections place the QQQ for sampling in the first row and second column of the quadrangle.

The NAPP photos corresponding with the selected quarter quad are then acquired. Finally, the appropriate quarter of the NAPP photo, corresponding to the randomly selected QQQ, is delineated as the area within which ground reference polygons will be defined.

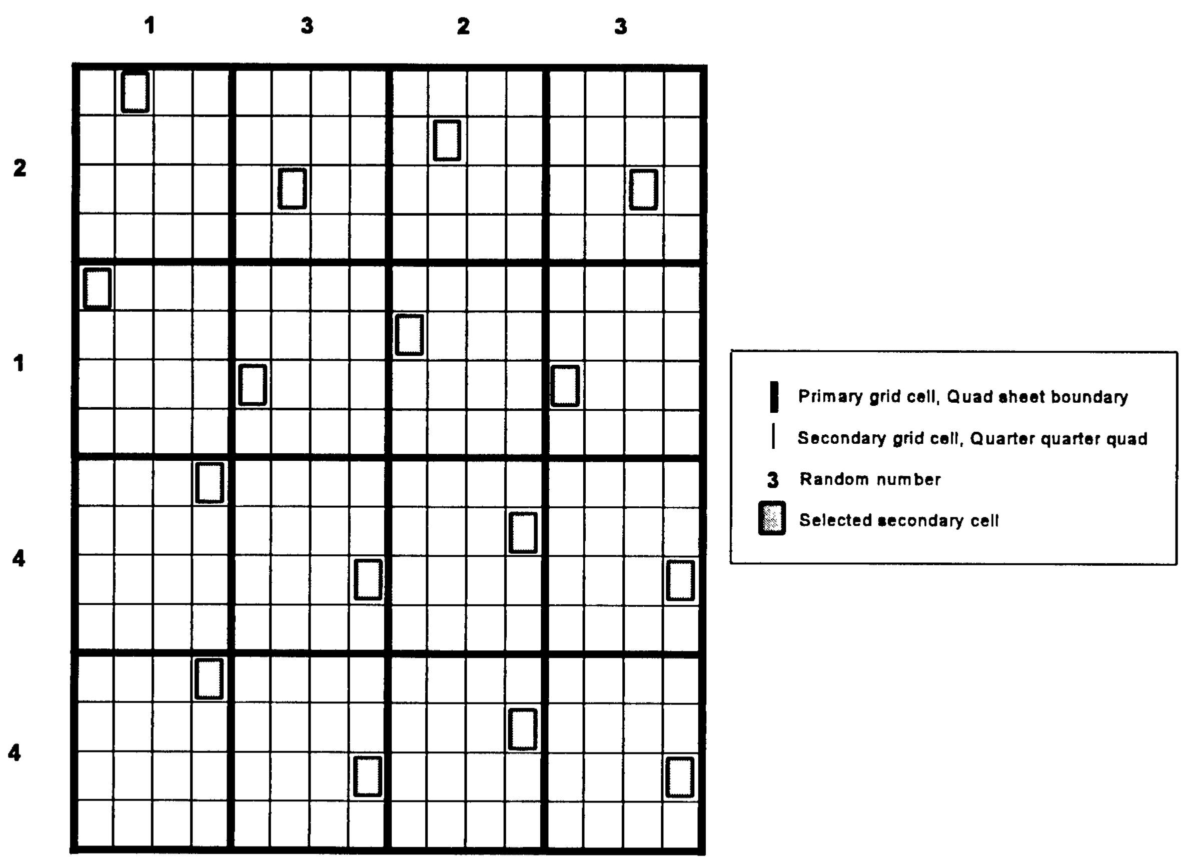


Figure 2. Geographically stratified sampling scheme with random eastings and northings, shown for 16 U.S. Geological Survey 7.5-min quadrangles.

3.2 Nonagricultural Sample Site Selection and Training

The NAPP photos selected using the above procedure are used by image analysts as a base for delineating ground reference data. It was determined that 9- x 9-inch contact prints at 1:40,000 scale would be adequate for this purpose. This format can be conveniently handled in the field and easily transported via mail.

In order to minimize staff time in the field and ensure that useful ground samples are collected, it was decided that sample sites should be chosen by image interpreters in the office, aided by viewing color composites of the TM data to be classified. First, a sheet of mylar is attached over each photo and the appropriate quarter of the NAPP photo is delineated. Next, image interpreters delineate candidate polygons on the mylar within the appropriate quarter using pencil. If sufficient auxiliary information is available to make an identification, the image analyst may pre-identify polygons to expedite the field-checking process.

Several criteria should be used when delineating polygons on photos. First, the polygons should be at least 2 ha. Second, the corresponding area on the TM imagery should be relatively homogenous in tone. Third, with few exceptions, the polygons should be delineated along roads. Fourth, the selected samples should be representative of the range of spectral variability present in the area, based on visual examination of the TM images. Following these guidelines will help ensure that each sample consists of only one cover type, that all cover types are sampled, and that staff can easily access the sites in the field (Figure 1).

As described above, it is important that the composition of the polygon set is representative of the variability in the stratum being used. Polygons may be delineated outside of the selected quarter photo when necessary to represent important spectral features not present in the selected quarter photo or when it is difficult to acquire a sufficient number of polyons in the selected quarter. It is also important to note that strata predominantly composed of agricultural cover will require fewer nonagricultural samples relative to the number of agricultural polygons.

Next, each polygon is assigned a unique number. The sample polygons are then delineated on the satellite imagery using screen digitizing to be used for future processing. The photos with mylar attached are delivered to field staff who field verify and record the UMGAP category associated with each ground sample polygon. Forms and definitions to be used by field staff are included in Appendix B.

Summary:	**Methods:**
1. Select the appropriate NAPP photo and position mylar overlay sheet.	1. Done manually.
2. Display the TM imagery for the corresponding area. Two images, three bands each, might be displayed side-by-side.	2. Display scenes in Viewer.
3. Select, number, and identify (if possible) at least 30 polygons, primarily within the selected quarter photo. Include polygons from other quarters of the photos as necessary. Polygons should be at least 2 ha and reasonably homogeneous in appearance in the raw TM data.	3. Done manually.
4. Delineate the selected polygons on the TM data, using screen digitizing.	4. Create vector coverage.
5. Deliver photos with mylar overlays to field personnel.	5. Done manually.

3.3 Agricultural Sample Site Selection and Training

The crop grown in any given field in the Upper Midwest may change annually (or even intra-annually) because of crop rotation. As a result, the collection date of agricultural ground reference data must match the TM acquisition date as closely as possible. To meet this requirement, photo bases and crop reports will be acquired from county Farm Service Agency (FSA) offices. These data are collected annually by FSA as part of that agency's 35-mm-based crop compliance program. Because these data are typically organized according to tracts of ownership, it is usually necessary to consult a plat map for each of the sections to be sampled to assist FSA in the information compilation process. That is, a list of owners by section usually must be compiled prior to making the information request to FSA.

Results of a pilot study at the Wisconsin Department of Natural Resources and the Environmental Remote Sensing Center showed that acquiring crop data for one public land survey section (nominally 1 mile × 1 mile) per QQQ is sufficient to provide agricultural training data for the agricultural base categories listed in Appendix A. The section chosen within the QQQ is deliberately selected by the image interpreter, based on the number of fields and diversity of crops within the section. It should be noted that more sections may be required in predominantly agricultural areas.

The boundary of each field is delineated on the imagery using screen digitizing. Some fields may be split into sub-samples to facilitate training and accuracy assessment.

3.4 Identification of Radiometric Normalization Reference Sites

One of the objectives of UMGAP is to provide useful data for land cover change-detection studies. There are a variety of different techniques used for change detection (Khorram et al. 1994; Lillesand and Kiefer 1994). Because some of these techniques require the radiometric standardization of multiple dates of imagery, it is important to be able to identify specific sites on the landscape that experience minimal spectral change over the anticipated period of change detection. These sites are used to radiometrically normalize one image to the other, in a process referred to as relative calibration. This approach was demonstrated by Coppin and Bauer (1994) in a multitemporal change-detection study in Minnesota and was recommended by the Coastal Change Assessment Program change-detection protocol (Khorram et al. 1994; Dobson et al. 1995).

Eckhardt et al. (1990) identified several important considerations for the selection of spectrally invariant sites used for radiometric normalization of multi-date images, including

- The sites must be of approximately the same elevation as the area of interest in the scene.
- The sites should contain little or no vegetation.
- The sites must be in a relatively flat area.
- When viewed on a display screen, the sites must have no apparent change in pattern over time.
- As far as possible, the sites should represent a wide range of pixel brightnesses.

During the UMGAP data collection and data processing stages, analysts should attempt to identify potential radiometric normalization sites. To the extent possible, from 10 to 20 well-distributed, radiometrically invariant sites should be identified in each scene. Ground targets will include such features as deep, nonturbid water bodies, roads, parking lots, rooftops, and other sites.

4. Satellite Image Data

Image data used for land cover classification can come from a variety of sensors, can be single date or multitemporal, and can be nearly raw or highly manipulated. This project is using two-date Landsat TM scenes, provided by the national GAP program (Jennings 1994). The multiple images that cover the project area need to be modified in several ways, including matching coordinate systems and eliminating areas of overlap between adjacent scenes.

4.1 Image Band Selection

The image band selection process was driven by two main criteria: the need for a high level of accuracy, and the need for efficient use of available computer resources. After a number of different tests, it was determined that the best results were obtainable using two-date TM imagery from all six reflectance (nonthermal) bands, compressed to three bands for each date by a principal components transformation. The TM imagery is well suited to this type of land cover classification because of its 30-m resolution and variety of spectral bands, especially in the near- and mid-infrared. The precise dates of imagery to be used vary from area to area as a result of both data availability and temporal variation in vegetation condition across the large area included in the study. In general, one TM image from summer and one from fall were selected to derive the most benefit from seasonal changes in forested areas. Spring and summer images were selected in areas dominated by agricultural cover types.

Because of the very large area involved, the processing and analysis of the 12 bands of data of the combined dates were considered to be a significant problem. Furthermore, it was anticipated that there would be a great deal of redundancy of information among the TM bands on each date because of interband correlation (Lillesand and Kiefer 1994). A number of studies have shown that principal components analysis (PCA) can be used to reduce the number of bands used in image analysis without significant loss of information (Jensen 1986). For this project, several different methods of generating the components were tried. The best results were achieved by creating separately the first three components from each date, then combining the two sets of components into a single six-band image for classification. Preliminary results showed that this combined principal components method produced as accurate classifications as did a larger number of raw image bands and involved significantly less time, effort, and disk space. To get the most benefit from the PCA process, any clouds present in the imagery are masked out prior to generating the principal component bands. Additionally, the principal components are generated separately for each stratum, rather than for the entire scene. These steps are described in more detail in Section 5.2, Scene Stratification.

4.2 Removal of Overlap for Adjacent Thematic Mapper Scenes

The numerous TM scenes that compose any state in the Upper Midwest overlap by approximately 35% on each side (and much less in the north-south direction). To reduce processing time, most of this overlap should be eliminated. Deciding which areas of overlap to eliminate is not trivial, especially in light of the need to further subdivide the states into spectrally consistent classification units (SCCUs), described in Section 5.2.

In the overlap area between two neighboring TM scenes, the image analyst must determine which portion of each image will be used for classification and which will be ignored. The two scenes can then be classified separately without processing the overlapping area twice. One consideration in eliminating overlap is the

presence of stratification unit boundaries (described in Section 5.2). Cloud cover, haze, and general image quality will also affect the decision of which portions of the overlapping areas to assign to a scene.

Screen digitizing is used to select the areas to be classified. A small amount of overlap (approximately 100 pixels) should remain between scenes. This area of overlap is used to compare the compatibility of the two classifications when completed and ensure that no gaps exist between images after they are stitched together.

5. The Classification Process

The UMGAP image processing methodology is the end-product of extensive research and development. It consists of two major procedures: stratification of the image data into several types of discrete units and classification of the pixels in each unit. These procedures are designed to maximize the accuracy and completeness of the resulting output maps. The entire process is described in proper order in a 14-step summary in Section 5.1.

Automated classification is the process of systematically extracting useful land cover information from raw remotely sensed imagery. The most well-developed methods of classification are based on analysis of spectral patterns among a set of image bands. A number of different classification algorithms have been employed; most such methods can be categorized as supervised, unsupervised, or hybrids of the two (Lillesand and Kiefer 1994). To determine the best automated classification methodology for this project, a series of tests was conducted and a set of protocols for the classification process was developed based on the results.

As described in Section 5.2, the satellite imagery are stratified in several ways. Where clouds are present, they are masked out. Next, urban areas are classified separately. Each scene is then broken up into a number of SCCUs, based in part on ecoregions but modified as necessary by photomorphic features of the imagery. Within each of these strata, wetlands are cut out (using existing digital wetlands boundary maps) and processed separately. The bulk of each stratum (the portion outside of all clouds, urban areas, and wetlands) is classified using a hybrid method referred to as guided clustering, followed by maximum likelihood classification. Wetlands are classified separately using traditional unsupervised clustering or guided clustering followed by maximum likelihood classification.

5.1 The Upper Midwest Gap Analysis Program Classification Process: A 14-Step Summary

The classification process consists of a series of 14 steps. These steps are described in more detail in Sections 5.2 through 5.6. To summarize the entire process, the 14 steps are listed here and are shown conceptually in Figure 3.

1. Delineate all cloud-covered areas in the scene and remove them from both image dates.

2. Delineate all urban areas and copy them from the parent images to separate files.

3. Compute principal components for urban areas separately for each date and combine the first three principal components from each date into a single urban principal component file.

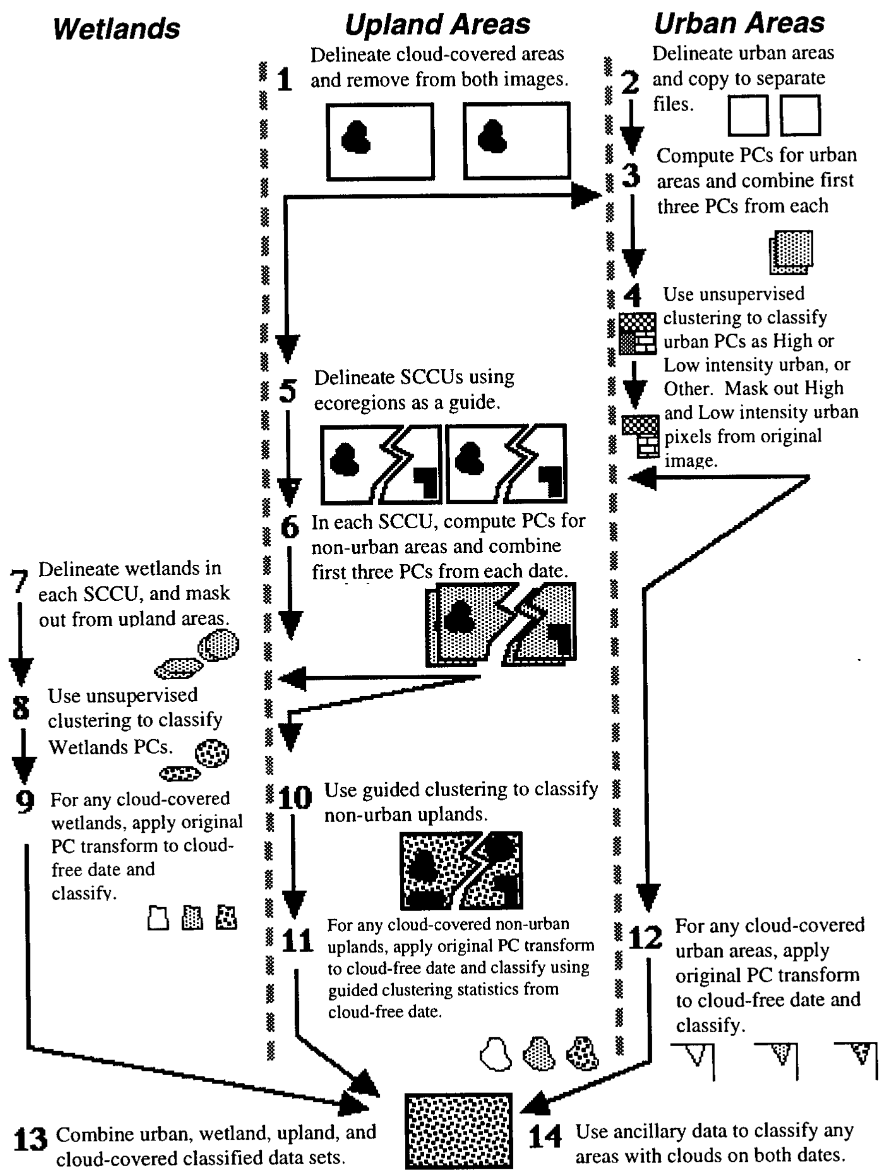


Figure 3. The Upper Midwest Gap Analysis Program classification process in 14 steps.

4. Use unsupervised clustering of the principal component bands to classify all urban areas into categories of "High intensity urban," "Low intensity urban," and "Other." Retain the "High intensity urban" and "Low intensity urban" pixels for subsequent replacement into the final classification and mask them out from the TM scenes. Do not retain "Other" pixels, which will be reclassified in the original image data set.

5. Delineate SCCUs in the original nonurban image data set based on photomorphic interpretation of the ecoregion map.

6. Within each SCCU, compute principal components for each image date separately for all remaining pixels in the parent data set (original - [clouds + "High intensity urban" + "Low intensity urban"]). Combine the first three principal components for each date into a single nonurban image data set.

7. Delineate all wetlands in each SCCU and remove them from the image.

8. Classify wetland areas in each SCCU using unsupervised clustering (or guided clustering) followed by maximum likelihood classification.

9. For any cloud-covered wetland areas, apply the original principal component transform to the cloud-free date and classify.

10. Classify nonurban upland areas in each SCCU using guided clustering followed by maximum likelihood classification.

11. For any cloud-covered nonurban uplands, apply the original principal component transform to the cloud-free date and classify using unsupervised clustering.

12. For any cloud-covered urban areas, apply the original principal component transform to the cloud-free date and classify.

13. Insert the "High intensity urban," "Low intensity urban," wetlands, and all single-date cloud-free classified areas into the nonurban upland classified data set.

14. Use ancillary data to classify all areas cloud covered in both image dates.

5.2 Scene Stratification

Classification projects in the past have realized improved accuracy as a result of scene stratification (Stewart 1994). This involves segmenting a large study area into smaller (more spectrally consistent) regions prior to classification. Several stratification methods were investigated for this project, including masking of urban areas, stratification by ecoregion, and subdivision of ecoregions using wetland/upland boundaries.

5.2.1 Clouds

If clouds are present in either date of imagery, screen digitizing is used to delineate them. The analyst visually identifies clouds in the imagery and also identifies cloud shadows based on their proximity to clouds. The clouds and cloud shadows are then masked out. During the classification process, these areas are classified based only on the data from the cloud-free date. Areas with clouds on both dates should be few in number and will either be classified using ancillary data only or left unclassified.

5.2.2 Urban Areas

Urban areas are often difficult to classify because they are a mixture of many cover types (Kramber and Morse 1994). Highly reflective urban cover is often confused with bare soil, resulting in errors of omission and commission with agriculture. Many authors have found that this problem can be overcome by classifying urban areas separately from nonurban areas (Robinson and Nagel 1990; Northcut 1991; Luman 1992).

Urban areas are copied to a separate file for classification. The TIGER Line Files from the 1990 Census are overlaid on an image backdrop as a guide and the analyst delineates boundaries around urban areas. The analyst may also refer to NAPP photos to assist in identifying urban areas. The urban areas are classified as high intensity urban, low intensity urban, or nonurban. After classification, those portions of the delineated urban areas classified as high intensity urban or low intensity urban are masked out of the TM images, whereas those portions of the delineated urban areas classified as nonurban are not masked out. Thus, any pixels within the delineated urban areas that have nonurban land cover will be classified with the remainder of the scene.

5.2.3 Spectrally Consistent Classification Units

Each scene is divided into several photomorphic SCCUs (Figure 4). These strata are based on ecoregion boundaries but are modified as necessary to delineate areas of relatively uniform appearance (including phenological regions and atmospheric influences) present in the image and not accounted for (or adequately represented) in the ecoregions. A variety of maps of ecoregions and landscape units have been proposed for stratification of remotely sensed data prior to classification (Stewart 1994); the SCCUs for UMGAP are based on the regional landscape ecosystems described by Albert (1995). After delineating SCCUs, the analyst should buffer each region by approximately 500 m, extending each into adjacent SCCUs, to assist in post-classification edge matching. At state borders, a buffer region extending approximately 3,000 m beyond the boundary should be included. As described in Section 4.1, principal components for each SCCU are generated separately for each date of imagery. The first three principal component bands from each date are then combined, making a single six-band image for each SCCU.

5.2.4 Wetlands

Numerous researchers have classified wetlands in the Upper Midwest with varied success (e.g., Best 1988; Cosentino 1992; Polzer 1992). Wetland classification accuracy is sometimes unacceptably low because wetland vegetation often appears spectrally similar to upland cover types. Because of this problem, it has been suggested that "current satellite technology is most valuable when used in conjunction with digital data derived from aerial photography and other sources" (Federal Geographic Data Committee 1992). For this reason, wetland surveys based on aerial photography, such as the National Wetlands Inventory, are being used to extract wetlands from each stratum of the satellite imagery after principal components are generated. Uplands and wetlands can then be processed separately. Only the most-generalized level of the wetlands inventory (wetlands versus uplands) is used to avoid tying the UMGAP classification to the potentially obsolete details of the photo-based inventory.

This procedure limits the confusion between upland and wetland types to those instances where errors of omission or commission exist in the wetlands inventory data. At the same time, using the satellite data for classification within wetland boundaries ensures that the classification of these areas is as current as possible and provides a uniform interpretation scale for both wetlands and uplands. For those who prefer the sometimes dated (but more detailed) National Wetlands Inventory data, these data can be "burned into" the TM classification at a later time.

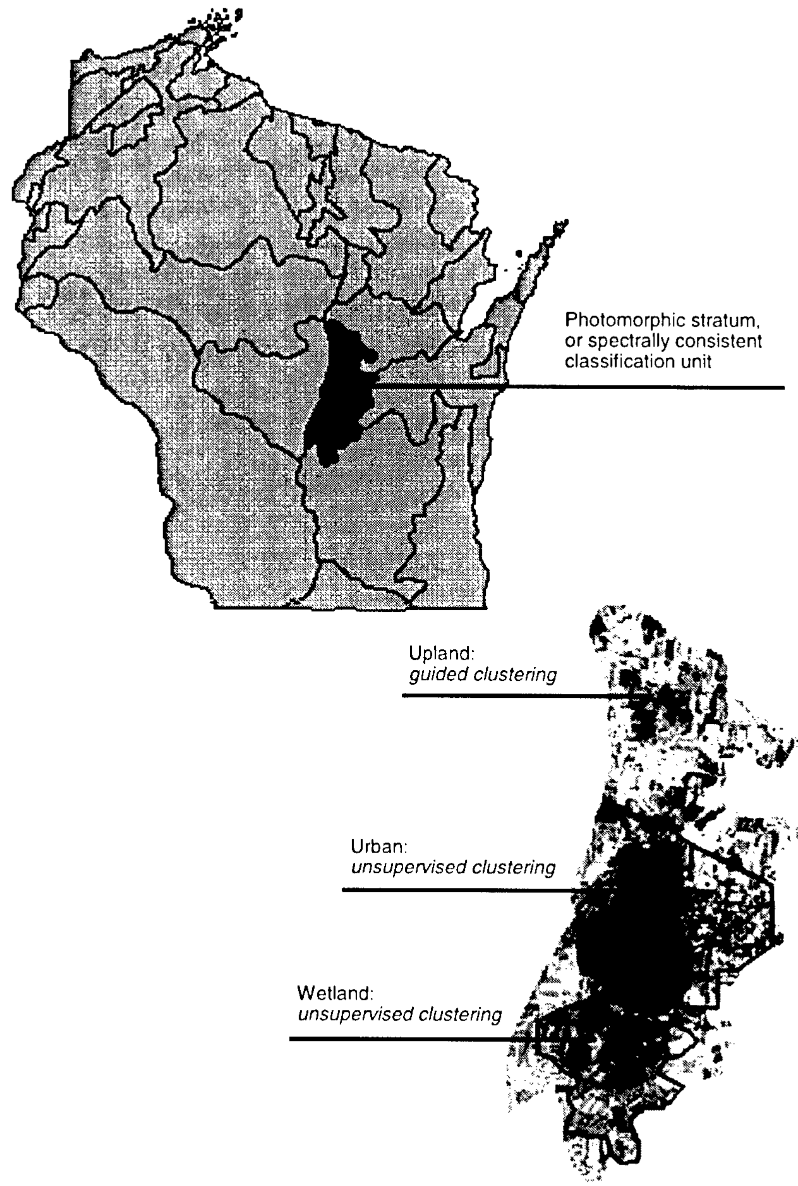


Figure 4. Preclassification image stratification.

Summary:

1. Use screen digitizing to delineate any clouds that appear on either date's image. Mask out these clouds.

2. Overlay TIGER Line files on the TM imagery and perform screen digitizing to delineate urban areas. Extract (copy) the urban areas from each date of

Methods:

1. Use "Mask" model (in-house) in Spatial Modeler.

2. Use AOI and Subset. For each date's image: Run Principal Components, in 16-bit mode, with the first three components for output. Run PCA

TM imagery, but do NOT mask them out. In the urban files, compute principal components separately for each date and combine the first three principal components from each date into a single file.	Stats Model (Imagine). Run C program (in-house) to format principal component statistics. Run principal component 16-to-8 bit adjustment model (in-house). Use Layer Stack to combine principal component files into a six-band file.
3. Classify the extracted urban area principal component bands into high intensity urban, low intensity urban, and nonurban classes. In the TM scene for each date, mask out pixels classified as high intensity urban or low intensity urban in the urban file. Do NOT mask out pixels within the delineated urban areas that were classified as nonurban.	3. See Section 5.3, Unsupervised Clustering of Urban Areas.
4. Overlay Albert's ecoregion boundaries on top of the image. Delineate SCCU boundaries, which reflect photomorphic features (including phenological regions and atmospheric influences) present in the image and are not accounted for, or accurately represented in, the ecoregions. A 500-m buffer should be left around the edge of each SCCU. Cut each date's image along the SCCU boundaries.	4. In Arc/Info, intersect ecoregions with outline of image to produce polygons. Build the new coverage. In Imagine, display image and overlay vectors. Use the Vector Query Tool to select polygons for AOI. Add selected polygons to AOI and save to file. Warp/Reshape AOIs to match photomorphic features. Use Subset with AOIs.
5. For each SCCU, generate principal component bands from the first date of imagery and from the second date of imagery. Combine the first three principal component bands from both images into a single file.	5. For each SCCU: Run Principal Components, in 16-bit mode, with the first three components for output. Run PCA Stats Model (Imagine), principal component stats formatting program (in-house), and principal component 16-to-8 bit adjustment model (in-house). Use Layer Stack to combine principal component files into a six-band file.
6. Import digitized wetland boundaries from photo-based inventory. Register the digitized wetland file to the TM imagery. Within each SCCU, overlay wetland polygons and extract wetland pixels. Set aside the wetlands portion for separate classification. Mask out the wetlands from the remaining (upland) portion of the SCCU.	6. In Imagine, display image and overlay vector wetlands file. Use Vector Query Tool to select polygons for AOI. Add selected polygons to AOI and save to file. Use Subset with AOIs. Use mask model (in-house) in Spatial Modeler to place 0s (zeros) in upland file.

5.3 Unsupervised Clustering of Urban Areas

When all of the urban areas have been delineated with screen digitizing, copy them from the TM imagery. Principal component bands are generated as described in Section 5.2. An unsupervised classification is performed on the extracted urban file, and the two urban classes, high intensity urban and low intensity urban, are differentiated. These pixels are masked out of the TM scene to be burned back in during the post-classification phase (see Section 6). All other pixels in the delineated urban areas are designated nonurban and are *not* masked out of the TM scene.

Because the urban areas were extracted prior to the creation of the SCCUs, all the urban areas in a scene are classified together.

Summary:	Methods:
1. Using an unsupervised ISODATA routine, cluster the extracted urban areas.	1. Using the AOIs from Section 5.2, run ISODATA with AOI option.
2. If desired, perform maximum likelihood classification of the urban areas with the clusters from ISODATA.	2. Run maximum likelihood classifier.
3. Recode subclasses as either high intensity urban, low intensity urban, or nonurban.	3. Use Recode.
4. Use the high intensity urban and low intensity urban pixels as a mask for the rest of the TM scene, as described in Section 5.2.	4. See Section 5.2.

5.4 Unsupervised Clustering of Wetlands

Wetland areas are cut from each SCCU during the stratification stage, after performing the principal components transformation described in Section 5.2 on each SCCU. The resulting wetlands-only portion of the TM image are clustered using an unsupervised ISODATA routine. Spectral clusters are labeled based on the wetlands inventory and other data sets as necessary. After classification of the remainder of the TM scene, the condensed wetland information classes are inserted into the final upland classification file. Note that extracting wetlands from the imagery should leave "holes" of zero value pixels in the TM data. This procedure should speed machine processing and mitigate confusion for image analysts concentrating on the upland data.

In some instances, when adequate training data are available, guided clustering may be used for wetlands classification rather than unsupervised clustering. The guided clustering methodology is described in Section 5.5.

Summary:	Methods:
1. Using an unsupervised ISODATA routine, cluster the wetlands-only portion of the TM image.	1. Using the AOIs from Section 5.2, run ISODATA with AOI option.
2. Perform maximum likelihood classification of the wetlands areas with selected clusters from ISODATA.	2. Run maximum likelihood classifier.
3. Label spectral clusters based on Wisconsin Wetlands Inventory or other data.	3. Recode classes.

5.5 Guided Clustering

Prior land cover classification projects have employed both supervised and unsupervised classification methods (Jensen 1986). Both methods, however, have inherent difficulties that make the classification process more costly and less reliable. Bauer et al. (1994) found that supervised techniques were inadequate for large-area classifications in the Upper Midwest region because of forest complexity, poor spectral separability, and the extensive manual processing required. In an attempt to resolve these problems with traditional supervised classification methods, a number of new techniques have been suggested.

Unsupervised techniques have the advantage of eliminating the costly and intensive training set delineation process of supervised classification, but identifying the resulting clusters can be difficult. Variability in different analysts' interpretation of the output of unsupervised classifiers may threaten the accuracy and objectivity of these classifications (McGwire 1992). Also, unsupervised classifiers reduce the ability of the analyst to control which classes are defined.

Guided clustering, the approach taken here, represents an alternative to supervised and unsupervised classification techniques (Lime and Bauer 1993; Bauer et al. 1994). It avoids most of the major pitfalls of the previous methods and appears well suited to large-area classifications with complex cover types. In guided clustering, the analyst delineates training sets for each cover type. Unlike the training sets used in traditional supervised clustering methods, these training sets need not be perfectly homogenous. For each information class, an unsupervised clustering routine is used to generate 20 or more spectral signatures from the class' training sets. These signatures are examined by the analyst; some may be discarded or merged and the remainder are considered to represent spectral subclasses of the desired information class. Signatures are also compared among the different information classes. Once a sufficient number of such spectral subclasses have been acquired for all information classes, a maximum likelihood classification is performed with the full set of refined spectral subclasses. The subclasses are then aggregated back into the original information classes.

Summary:	**Methods:**
1. The analyst delineates training pixels for information class X.	1. Use Vector Query Tool with Arc coverage. Use query to select polygons based on SCCU ID, class, and assessment or training status. Convert to AOI.
2. Cluster class X pixels into spectral subclasses X1..Xn using an automated clustering algorithm.	2. ISODATA.
3. Examine class X signatures and merge or delete signatures as appropriate. A progression of clustering scenarios (e.g., from 3 to 20) should be investigated, with the final number of clusters and merger and deletion decisions based on such factors as (1) display of a given class on the raw image, (2) multidimensional histogram analysis for each cluster, and (3) multivariate distance measures (e.g., transformed divergence or Jeffries-Matusita distance).	3. Evaluate signatures in Signature Editor and modify as desired.
4. Repeat steps 1–3 for all additional information classes.	4. Repeat steps 1–3. Use Append option in Signature Editor to unite all spectral signatures for all classes in a single file.
5. Examine ALL class signatures and merge or delete signatures as appropriate.	5. Evaluate signatures in Signature Editor and modify as desired.
6. Perform maximum likelihood classification on the entire SCCU with the full set of spectral subclasses, saving the Probability Density Function image.	6. Run maximum likelihood classifier.
7. Aggregate spectral subclasses back to the original information classes.	7. Use Recode.

To ensure that all of the spectral classes present in a SCCU are represented, the analyst may perform an unsupervised clustering of the entire SCCU as a test. The resulting cluster signatures are compared to the full set of spectral signatures from guided clustering to help determine whether any significant spectral classes have been omitted. If the unsupervised clustering produces any clusters that are not well represented by any of the signatures developed through guided clustering, additional training samples may be required.

If any clouds were present in a particular SCCU, the clouded areas masked out in Section 5.2 will have to be classified in a separate step after the rest of the SCCU is classified. The same set of signatures created during the guided clustering of the noncloudy portion of the SCCU will still be used for the cloud covered areas. However, the signature files must be edited to remove the three principal component bands for the cloudy image. The maximum likelihood classification will then be done using only the bands from the cloud-free image.

5.6 Maximum Likelihood Classification

Statistical classifiers in image processing have proven successful in many land cover classification projects. In general, these classifiers assign an image pixel to its most likely class, based upon the class mean, variance, and covariance in each band. This process may involve calculating a number of different probability values representing the likelihood that a given pixel belongs to each of the spectral classes in the final classification. For some applications, it may be desirable to have an indication of the likelihood that a given pixel is actually a member of the class to which it was assigned. For this reason, the maximum likelihood classifier will save an image of the probability density function from each classification. These images will aid in identifying areas and classes of questionable accuracy. The probability density function images,for each stratum are used interactively during the classification process. They are also saved for future reference by users who wish to have access to information about the spatial variability and class variability of the classification probabilities.

5.7 Alternative Classification Methods

The classification methods described here are designed to be standardized and repeatable and to permit replication elsewhere under varying conditions. For some portions of the tristate Upper Midwest Gap Analysis Project, however, it may be desirable to consider alternative classification strategies. One example of such an alternative strategy is the use of carefully timed multiseason imagery designed to maximize the benefit of phenological variability (e.g., Wolter et al. 1995). Before deciding on an alternative classification method, it is important to carefully examine the nature of the proposed classification strategy and to determine whether it satisfies all of the design considerations presented in this document.

6. Post-Classification Processing

As each scene is classified to an acceptable level of accuracy, it can be used to aid in classifying neighboring images. When an initial classification is completed for any given SCCU, it should be compared to all of its neighbors whose accuracy has already been assessed. Distinct differences along the boundary between the two scenes could indicate that the classification in question will need modifications. This process will help mitigate categorical edge-matching errors when the scenes or strata are finally stitched together.

After each SCCU has been classified, the wetlands, urban areas, and cloud-covered pixels extracted from it and separately classified are placed back in the image. Transportation features, such as roads and railroads, are then added into the classified image from ancillary sources such as USGS Digital Line Graphs. A variety of products will be generated from the classified imagery. Digital versions of the data will be made available in both raw and filtered formats, to meet the needs of different end users. For filtered products, a clump-and-sieve algorithm is used. Adjacent pixels sharing the same class are grouped into clumps. Clumps smaller than four pixels in size are deleted and the resulting holes are filled in by expansion of neighboring clumps. The clump-and-sieve process is performed separately on upland and wetland areas to prevent upland areas from extending into wetlands and vice versa. In addition, pixels classified as water are preserved regardless of clump size. Note that for filtered data, the probability density function images produced during maximum likelihood classification will not be applicable. In addition to digital data, hard-copy products can be generated at a variety of scales. Finally, to meet the national GAP project standards, the data will also be "vectorized" (converted to vector format) and aggregated to a 100-/40-ha minimum mapping unit at the Environmental Management Technical Center (Jennings 1994).

Summary:	**Methods:**
1. Add any delineated areas with clouds back into the SCCU from which they were originally extracted.	1. Use Class Merge Model (Spatial Modeler), with clouds and full scene. If <raster> <> 0 use <raster>.
2. Add the classified wetlands pixels back into the SCCU from which they were originally extracted.	2. Use Class Merge Model (Spatial Modeler), with wetlands and full scene. If <raster> <> 0 use <raster>.
3. Stitch together neighboring SCCUs, examining boundaries for discontinuities.	3. Use Subset.
4. Add the classified urban area pixels back into the classified scene.	4. Use Class Merge Model (Spatial Modeler), with urban areas and full scene. Select only "high intensity urban" and "low intensity urban" to be placed back in the full scene.
5. Overlay transportation features from USGS Digital Line Graph files on top of the classified image.	5. Vector Overlay.

7. Accuracy Assessment

Few aspects of the land cover mapping process are as elusive and challenging as assessing the accuracy of the final products resulting from such efforts. The literature includes several recent treatises specifically focused on the subjects of classification accuracy assessment (e.g., Congalton 1991; Janssen and van der Wel 1994) and land cover change-detection accuracy assessment (e.g., Khorram et al. 1994). These documents highlight the need to consider both the positional accuracy and thematic accuracy of any given data product.

7.1 Positional Accuracy Considerations

The data used for UMGAP classification have been registered to the Universal Transverse Mercator coordinate system (e.g., Universal Transverse Mercator or Wisconsin Transverse Mercator) and subsequently resampled (primarily using cubic convolution). Through the careful selection of numerous, well-defined, and well-distributed ground control points (GCPs), the positional accuracy (RMSE) of well-defined objects appearing in the TM imagery should be on the order of ± 0.5 pixels, or ± 15 m. Also, registration of one

TM scene to another is expected to be on the order of ± 0.5 pixels and no more than ± 1 pixel. Ideally, the georeferencing of each scene should be verified using a minimum of 10 GCPs (with a minimum of 2 GCPs in each quadrant of the scene) and 7.5-min quadrangles. Care should be taken to ensure that the same datum (e.g., NAD83) is used for the check as was used for the original scene georeferencing process. Scenes with RMSE values in excess of ± 1 pixel should be reregistered.

7.2 Thematic Accuracy Considerations

7.2.1 Anticipation of Multipurpose Use of Upper Midwest Gap Analysis Program Land Cover Data

It is anticipated that UMGAP land cover data will be used over a range of geographic scales from the site to the statewide level. No single thematic accuracy assessment methodology is appropriate over this range of applications. Accordingly, the philosophy of the thematic accuracy assessment protocol for UMGAP is to provide sufficient raw information at a base level to enable a flexible range of potential accuracy assessment scenarios in various future application contexts. The following information relates to the collection of base level data only.

7.2.2 Sample Unit

The fundamental sample unit available for accuracy assessment is the polygon, for this is the unit within which the ground reference data are collected. A census of all pixels in the polygon is performed to determine the most abundant class within the polygon. In most cases, a single class should be clearly dominant because the ground reference data collection effort in which the polygons were delineated was designed to include only homogenous areas. The analyst should visually examine accuracy assessment polygons to ensure that this is the case.

7.2.3 Reference Data for Accuracy Assessment

Section 3, Ground Reference Data, describes some of the methods used for collecting reference data for UMGAP. The methods used are not completely random because of the focus on rapid and cost-effective acquisition of a large volume of representative data for training purposes. Only a portion of the data collected are required for training, and the remainder can be used to help assess the accuracy of the final classifications. It is important to note, however, that many of the statistical techniques described below are based upon an assumption of randomness. In particular, the fact that reference polygons are selected and delineated manually results in unequal (and unknowable) probabilities of inclusion for different points on the ground. This may introduce a bias into the estimators for categorical and overall accuracy and may also affect the estimators for the variance of these quantities (Czaplewski 1994). Future investigations are planned to evaluate the effectiveness of data collection methods for a variety of accuracy assessment strategies.

7.2.4 Classification Error Matrices

The most widely used accuracy assessment techniques for land cover classification involve the use of error matrices as the primary basis for comparing, on a category-by-category basis, the relation between the known reference data (columns) and the corresponding results of the automated classification (rows). In

addition to compilation of the complete matrix, the following descriptive statistics can be computed: overall accuracy, producer accuracy of each category, user accuracy of each category, the two-tailed 95% confidence interval of the overall accuracy and the producer and user accuracies, and the Kappa (KHAT) statistic for the overall classification and each individual category (Lillesand and Kiefer 1994). Examples of the computation of these descriptive statistics are contained in Appendix C.

7.3 Other Accuracy Assessment Products

Certain specialized accuracy assessment products will be available from the UMGAP classification process. These include storage and cartographic portrayal of the probability density function value associated with the most probable class assignment of each pixel by the maximum likelihood algorithm. Also, the integration of the accuracy assessment and training sampling process permits depiction of the exact areas used for accuracy assessment. The polygons used for this process are stored in a vector file that is automatically registered to the same coordinate system as the image data. Thus, it is possible to document the distribution of accuracy assessment sites by overlaying this vector file directly on the raw imagery, on a USGS topographical map, or another georeferenced data source.

8. Conclusion

This document was written to explain and codify the image processing procedures in the UMGAP land cover classification being performed with multi-date TM data. These procedures continue to evolve as they are employed in a production environment. Also, they are intended to be the basis for the initial land cover classification involved in UMGAP. New data sources and methods continually enhance the approaches described herein. Our objective was to provide a firm foundation for these anticipated enhancements.

9. Acknowledgments

Numerous individuals and agencies have participated in the production of this document. The form of their involvement has ranged from actual writing of various sections, to critical review of preliminary drafts, to providing substantive input during numerous meetings held on the subject of the protocol, to funding the preliminary and continuing research on which the protocol is based. Space precludes our specific identification of all of these individuals and agencies.

Much of the protocol results from the collective effort of personnel from the University of Wisconsin–Madison Institute for Environmental Studies Environmental Remote Sensing Center working closely with members of the staff of the Wisconsin Department of Natural Resources. Contributors to this collective effort include Jana Stewart, who performed the background research leading to the image stratification methods specified in the protocol, and Thomas Simmons and Thomas Ruzycki, who provided input to the protocol's development. On behalf of Wisconsin Department of Natural Resources, Paul Tessar was responsible for engendering the agency's role in the formation and implementation of the WISCLAND. Robert F. Gurda, Assistant State Cartographer, is recognized for his invaluable role in chairing the WISCLAND interagency steering committee.

Many aspects of this protocol were influenced by various and numerous contributions made by personnel from both the University of Minnesota Remote Sensing Laboratory and the Minnesota Department of Natural Resources. Marvin E. Bauer and his colleagues at the Remote Sensing Laboratory performed the preliminary

research leading to the adaptation of the hybrid guided clustering procedures specified in the protocol. Much of this work was performed in cooperation with Minnesota Department of Natural Resources staff, including William Befort and David F. Heinzen. They, and several other Minnesota Department of Natural Resources staff, have played a very active and important role in developing the protocol and implementing Gap analysis in Minnesota.

Dale Rabe and Michael Donovan of the Michigan Department of Natural Resources have been primarily responsible for implementing the image processing aspects of the Gap analysis being conducted in the state of Michigan. This effort is being conducted in close cooperation with Peter Joria of the USGS, Environmental Management Technical Center.

The Environmental Management Technical Center has been responsible for the overall coordination of the entire UMGAP. Frank D'Erchia, as UMGAP Principal Investigator, and Daniel Fitzpatrick, as Biodiversity Coordinator, are particularly acknowledged for their roles in providing the administrative and technical "glue" to hold such a complex tristate effort together and moving in a coherent direction.

References

Albert, D. A. 1995. Regional landscape ecosystems of Michigan, Minnesota, and Wisconsin. General Technical Report NC-178. North Central Forest Experiment Station, U.S. Forest Service, St. Paul, Minnesota. 250 pp.

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer. 1976. A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964. 28 pp.

Bauer, M. E., T. E. Burk, A. R. Ek, P. R. Coppin, S. D. Lime, T. A. Walsh, D. K. Walters, W. Befort, and D. F. Heinzen. 1994. Satellite inventory of Minnesota forest resources. Photogrammetric Engineering and Remote Sensing 60(3):287–298.

Berry, B. J. L., and A. M. Baker. 1968. Geographic sampling. Pages 91–100 *in* B. J. L. Berry and D. F. Marble, editors. Spatial analysis—A reader in statistical geography. Prentice-Hall, Englewood Cliffs, New Jersey.

Best, R. G. 1988. Use of satellite data for monitoring parameters related to the food habits and physical condition of Canada geese in Wisconsin during spring migration. Ph.D. Thesis, University of Wisconsin–Madison. n.p.

Congalton, R. G. 1988. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54(5):593–600.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of the Environment 37:35–46.

Congalton, R. G., and R. A. Mead. 1983. A quantitative method to test for consistency and correctness in photointerpretation. Photogrammetric Engineering and Remote Sensing 49(1):69–74.

Coppin, P., and M. E. Bauer. 1994. Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features. IEEE Transactions on Geoscience and Remote Sensing 32(4):918–927.

Cosentino, B. L. 1992. Satellite remote sensing techniques in support of natural resource monitoring: A view towards statewide land cover mapping. Masters Thesis, University of Wisconsin-Madison. n.p.

Czaplewski, R. L. 1994. Variance approximations for assessments of classification accuracy. U.S. Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado, Research Paper RM-316. 29 pp.

Dobson, J. E., E. A. Bright, R. L. Ferguson, D. W. Field, L. L. Wood, K. D. Haddad, H. Iredale III, J. R. Jensen, V. V. Klemas, R. J. Orth, and J. P. Thomas. 1995. NOAA Coastal Change Analysis Program (C-CAP): Guidance for Regional implementation. U.S. Department of Commerce, Seattle, Washington, NOAA Technical Report NMFS 123. 92 pp.

Eckhardt, D. W., J. P. Verdin, and G. R. Lyford. 1990. Automated update of an irrigated lands GIS using SPOT HRV imagery. Photogrammetric Engineering and Remote Sensing 56(11):1515–1522.

Faber-Langendoen, D. 1993. Midwest regional community classification. The Nature Conservancy, Midwest Regional Office, Minneapolis, Minnesota. 22 pp.

Federal Geographic Data Committee-Wetlands Subcommittee. 1992. Application of satellite data for mapping and monitoring wetlands. Technical Report 1, Washington, D.C. n.p.

Hudson, W. D., and C. W. Ramm. 1987. Correct formulation of the Kappa coefficient of agreement. Photogrammetric Engineering and Remote Sensing 53(4):421–422.

Janssen, L. L. F., and F. J. M. van der Wel. 1994. Accuracy assessment of satellite-derived land-cover data: A review. Photogrammetric Engineering and Remote Sensing 60(4):419–426.

Jennings, M. D. 1994. National Gap Analysis Project Standards, revision of November 1994. Gap Analysis Project, National Biological Survey and Idaho Cooperative Fish and Wildlife Research Unit, Moscow, Idaho. n.p.

Jensen, J. R. 1986. Introductory digital image processing: A remote sensing perspective. Prentice-Hall, Inc., Englewood Cliffs, New Jersey. 379 pp.

Khorram, S., G. S. Biging, N. R. Chrisman, D. R. Colby, R. G. Congalton, J. E. Dobson, R. L. Ferguson, M. F. Goodchild, J. R. Jensen, and T. H. Mace. 1994. Accuracy assessment of land cover change detection. North Carolina State University, Raleigh, Computer Graphics Center Report 101. 70 pp.

Klemas, V. V., J. E. Dobson, R. L. Ferguson, and K. D. Haddad. 1993. A coastal land cover classification system for the NOAA Coastwatch Change Analysis Project. Journal of Coastal Research 9(3):862–872.

Klemas, V. V., S. R. Hoffer, R. Kleckner, D. Norton, and B. O. Wilen. 1992. A modified wetland/upland land cover classification system for use with remote sensors. Pages 65–69 *in* U.S. Geological Survey, Reston, Virginia, Forum on Land Use and Land Cover Summary Report.

Kramber, W. J., and A. Morse. 1994. Integrating image interpretation and unsupervised classification procedures. 1994 ASPRS/ACSM Annual Convention and Exposition Technical Papers, Reno, Nevada 1:327–336.

Lillesand, T. M. 1994. Strategies for improving the accuracy and specificity of large-area, satellite-based land cover inventories. *In* Proceedings, ISPRS Mapping and GIS Symposium, Athens, Georgia 30:23–30.

Lillesand, T. M., and R. W. Kiefer. 1994. Remote sensing and image interpretation, 3rd edition. Wiley, New York. 750 pp.

Lime, S. D., and M. E. Bauer. 1993. Guided clustering. University of Minnesota Remote Sensing Laboratory Technical Memorandum. n.p.

Luman, D. E. 1992. Lake Michigan Ozone study final report. Northern Illinois University, Department of Geography and Center for Governmental Studies. 58 pp.

McGwire, K. C. 1992. Analyst variability in labeling of unsupervised classifications. Photogrammetric Engineering and Remote Sensing 58(12):1673–1677.

Northcut, P. 1991. The incorporation of ancillary data in the classification of remotely sensed data. Masters Thesis, University of Wisconsin–Madison. n.p.

Ott, L. 1988. An introduction to statistical methods and data analysis, 3rd edition. PWS-Kent, Boston. 835 pp.

Polzer, P. 1992. Assessment of classification accuracy improvement using multispectal satellite data: Case study in the glacial habitat restoration area of east central Wisconsin. Masters Thesis, University of Wisconsin-Madison. 110 pp.

Robinson, R., and D. Nagel. 1990. Land cover classification of remotely sensed imagery and conversion to a vector-based GIS for the Suwannee River water management district. Pages 219–224 *in* Proceedings: 1990 GIS/LIS, Anaheim, California.

Rosenfield, G. H., and K. Fitzpatrick-Lins. 1986. A coefficient of agreement as a measure of thematic classification accuracy. Photogrammetric Engineering and Remote Sensing 52(2):223–227.

Scott, J. M., F. Davis, B. Csuti, R. Noss, B. Buterfield, C. Groves, H. Anderson, S. Caicco, F. D'Erchia, T. C. Edwards, Jr., J. Ulliman, and R. G. Wright. 1993. Gap analysis: A geographic approach to protection of biological diversity. Wildlife Monograph 123:1–41.

Snedecor, G. W., and W. G. Cochran. 1989. Statistical methods, 8th edition. Iowa State University Press, Ames.

Stewart, J. S. 1994. Assessment of alternative methods for stratifying Landsat TM data to improve land cover classification accuracy across areas with physiographic variation. Masters Thesis, University of Wisconsin–Madison. n.p.

Thomas, I. L., and G. M. Allcock. 1984. Determining the confidence level for a classification. Photogrammetric Engineering and Remote Sensing 50(10):1491–1496.

United Nations Educational, Scientific, and Cultural Organization. 1973. International classification and mapping of vegetation. United Nations Educational, Scientific, and Cultural Organization, Paris. 35 pp.

Wolter, P. T., D. J. Mladenoff, G. E. Host, and T. R. Crow. 1995. Improved forest classification in the Northern Lake States using multi-temporal landsat imagery. Photogrammetric Engineering and Remote Sensing 61(9):1129–1143.

Appendix A. Upper Midwest Gap Analysis Program Classification System

Base categories are in **boldface.** *Extended categories are in* plain text. *Eight-bit numeric ID numbers are listed in parentheses (). * denotes classes limited to Minnesota. ‡ denotes classes limited to Wisconsin.*

- **(100) 1 Urban/developed**
  - **(101) 1.1 High intensity**
  - **(104) 1.2 Low intensity**
  - **(107) 1.3 Transportation**

- **(110) 2 Agriculture**
  - **(111) 2.1 Herbaceous/field crops**
    - **(112) 2.1.1 Row crops**
      - **(113) 2.1.1.1 Corn**
      - (114) 2.1.1.2 Peas ‡
      - (115) 2.1.1.3 Potatoes ‡
      - (116) 2.1.1.4 Snap beans ‡
      - (117) 2.1.1.5 Soybeans ‡
      - **(118) 2.1.1.6 Other**
    - **(124) 2.1.2 Forage crops**
      - (125) 2.1.2.1 Alfalfa ‡
    - (131) 2.1.3 Small grain crops ‡
      - (132) 2.1.3.1 Oats ‡
      - (133) 2.1.3.2 Wheat ‡
      - (134) 2.1.3.3 Barley ‡
  - (140) 2.2 Woody
    - (141) 2.2.1 Nursery
    - (144) 2.2.2 Orchard
    - (147) 2.2.3 Vineyard

- **(150) 3 Grassland**
  - (151) 3.1 Cool season
  - (154) 3.2 Warm season
  - (157) 3.3 Old field

- **(160) 4 Forest**
  - **(161) 4.1 Coniferous**
    - **(162) 4.1.1 Jack pine**
    - **(163) 4.1.2 Red/white pine**
    - (164) 4.1.3 Scotch pine ‡
    - (165) 4.1.4 Hemlock ‡
    - (166) 4.1.5 White spruce
    - (167) 4.1.6 Norway spruce ‡
    - (168) 4.1.7 Balsam fir
    - (169) 4.1.8 Northern white-cedar
    - **(173) 4.1.9 Mixed/other coniferous**
  - **(175) 4.2 Broad-leaved deciduous**
    - **(176) 4.2.1 Aspen**
    - **(177) 4.2.2 Oak**
      - (178) 4.2.2.1 White oak
      - (179) 4.2.2.2 Northern pin oak
      - (180) 4.2.2.3 Red oak
    - (181) 4.2.3 White birch
    - (182) 4.2.4 Beech ‡
    - (183) 4.2.5 Maple
      - (184) 4.2.5.1 Red maple
      - (185) 4.2.5.2 Sugar maple
    - (186) 4.2.6 Balsam poplar *
    - **(187) 4.2.7 Mixed/other broad-leaved deciduous**

	(190)	**4.3**	**Mixed deciduous/coniferous**		
		(191)	4.3.1	Pine-deciduous *	
			(192)	4.3.1.1	Jack pine-deciduous *
			(193)	4.3.1.2	Red/white pine-deciduous *
		(194)	4.3.2	Spruce/fir-deciduous *	
(200)	**5**	**Open water**			
(210)	**6**	**Wetland**			
	(211)	**6.1**	**Emergent/wet meadow**		
		(212)	6.1.1	Floating aquatic *	
		(213)	6.1.2	Fine-leaf sedge *	
		(214)	6.1.3	Broad-leaved sedge-grass *	
		(215)	6.1.4	Sphagnum moss *	
	(217)	**6.2**	**Lowland shrub**		
		(218)	6.2.1	Broad-leaved deciduous	
		(219)	6.2.2	Broad-leaved evergreen	
		(220)	6.2.3	Needle-leaved	
	(222)	**6.3**	**Forested**		
		(223)	**6.3.1**	**Broad-leaved deciduous**	
			(224)	6.3.1.1	Red maple
			(225)	6.3.1.2	Silver maple *
			(226)	6.3.1.3	Black ash
			(227)	6.3.1.4	Mixed/other deciduous *
		(229)	**6.3.2 Coniferous**		
			(230)	6.3.2.1	Black spruce
			(231)	6.3.2.2	Tamarack
			(232)	6.3.2.3	Northern white-cedar
		(234)	**6.3.3 Mixed deciduous/coniferous**		
(240)	**7**	**Barren**			
	(241)	7.1	Sand		
	(242)	7.2	Bare soil		
	(245)	7.3	Exposed rock		
	(246)	7.4	Mixed		
(250)	**8**	**Shrubland**			

Appendix B. Sample Ground Reference Data Forms and Definitions

Please check the land cover type associated with the polygon-ID. Choose that which best describes the land cover; land cover type definitions are provided on an enclosed sheet. **Please read the definitions prior to groundtruthing**. Record additional comments, such as species information for nonforested cover types, or percent composition for mixed categories, such as shrub and grassland, in the comments section.

NAME: **DATE:**
NAPP PHOTO-ID: **POLYGON-ID:**

(1) COVER TYPE

URBAN/DEVELOPED
____ High Intensity Urban
____ Low Intensity Urban

AGRICULTURE
____ Row Crops
____ Forage Crops

FOREST
____ Coniferous
____ Broad-leaved Deciduous
____ Mixed Coniferous/Broad-leaved Deciduous
____ Clearcut/Young Plantation - If clearcut, was area logged within the past 3 years? Circle: Yes or No

SHRUBLAND
____Upland Shrub

GRASSLAND
____Grassland

OPEN WATER
____ Open Water

BARREN
____Sand
____Bare Soil
____Exposed Rock
____Mixed

WETLAND
____Emergent/Wet Meadow
____Lowland Shrub
____Coniferous
____Broad-leaved Deciduous
____Broad-leaved Evergreen
____Forested Wetland
____Coniferous
____Broad-leaved Deciduous
____Mixed Coniferous/ Broad-leaved Deciduous

Comments: ____________________

(2) FOREST SPECIES

Write the estimated percentage of the species present in the space provided.
The percentages should total the canopy cover percentage in section 3.

____ % Jack Pine
____ % Red Pine
____ % White Pine
____ % Black Spruce
____ % White Spruce
____ % Balsam Fir
____ % Hemlock
____ % White Cedar
____ % Tamarack
____ % Aspen

____ % Red Maple
____ % Sugar Maple
____ % Silver Maple
____ % Green Ash
____ % Black Ash
____ % White Birch
____ % Yellow Birch
____ % River Birch
____ % Basswood

____ % Alder
____ % Red/Black Oak
____ % White/Bur Oak
____ % N. Pin Oak
____ % Slippery Elm
____ % Amer. Elm
____ % Black Cherry

____ % Black Willow
____ % Cottonwood
____ % Beech

Other Species
____ % ____________
____ % ____________
____ % ____________

Are trees at mature height? Circle: Yes or No
Comments: ____________________

(3) CANOPY AND UNDERSTORY

Canopy cover is: _____%

If canopy is less than 80%, mark the understory vegetation present:
____ Small trees
____ Saplings
____ Shrubs
____ Herbaceous Vegetation

Comments: ____________________

(4) METHOD OF IDENTIFICATION
____ Field Verification (Able to identify location and access the area circled.)
____ Windshield Survey (Could not enter identified area, but identified species from outside of area.)
____ Inaccessible Polygon
____ Photo interpreted / Knowledge of area

(5) CONFIDENCE LEVEL OF ASSESSMENT
______ High (good) ______ Medium _____ Low (questionable)

(6) ADDITIONAL COMMENTS

____

Definitions to Accompany Groundtruth Data Sheets

I. URBAN/DEVELOPED

Structures and areas associated with intensive land use.

a. **High Intensity** - Greater than 50% solid impervious cover of synthetic materials.

Examples: parking lot, shopping mall, or industrial park

b. **Low Intensity** - Less than 50% solid impervious cover of synthetic materials. May have some interspersed vegetation.

Examples: sparse development, single family residence

Note: Areas meeting the requirements of both Urban/Developed and Forest classes should be classified in the Urban/Developed category. (i.e., residential areas with greater than 10% crown closure of trees would be classified as Urban/Developed, rather than forest.)

II. AGRICULTURE

Land under cultivation for food or fiber (including bare or harvested fields).

Examples: corn, peas, alfalfa, wheat, orchards, cranberry bogs

III. GRASSLAND

Lands covered by noncultivated herbaceous vegetation predominated by grasses, grass-like plants or forbs.

Examples: cool or warm season grasses, restored prairie, abandoned fields, golf course, sod farm, hay fields

IV. FOREST

An upland area of land covered with woody perennial plants, the tree reaching a mature height of at least 6 feet tall with a definite crown. Crown closure of the area must be greater than 10%.

a. **Coniferous** - Upland areas whose canopies have a predominance (greater than 33-1/3%) of cone-bearing trees, reaching a mature height of at least 6 feet tall. If the deciduous species group is present, it should not exceed one-third (33-1/3%) of the canopy.

Examples: Jack Pine, Red Pine, White Spruce, Hemlock, Tamarack

b. **Broad-leaved Deciduous** - Upland areas whose canopies have a predominance (greater than 33-1/3%) of trees, reaching a mature height of at least 6 feet tall, which lose their leaves seasonally. If the coniferous species group is present, it should not exceed one-third (33-1/3%) of the canopy.

Examples: Aspen, Oak, Maple, Birch

c. **Mixed Coniferous/Broad-leaved Deciduous** - Upland areas where deciduous and evergreen trees are mixed so that neither species **group** (broad-leaved deciduous or coniferous) is less than one-third (33-1/3%) dominant in the canopy.

Examples: Hemlock/Northern Hardwood forest (40% Coniferous, 60% Broad-leaved Deciduous)

d. **Clearcut/Young Plantation** - Area used for tree production that has been recently cut, and is generally devoid of established vegetation cover, with the continued intention of tree production. Also an area that has been very recently replanted with trees (usually as a monoculture). *If the area has been logged within the last 3 years, please indicate this in the comments section of the groundtruth sheet.*

Note: Areas that meet the requirements of both Forest and Forested Wetland categories should be classified in the Forested Wetland category.

V. OPEN WATER

Areas of water with no vegetation present.

Examples: Lake, Reservoir, River, Retaining Pond

VI. WETLAND

An area with water at, near, or above the land surface long enough to be capable of supporting aquatic or hydrophytic vegetation, and with soils indicative of wet conditions.

a. **Emergent/Wet Meadows** - Persistent and nonpersistent herbaceous plants standing above the surface of the water or soil.

Examples: Cattails, Marsh Grass, Sedges

b. **Lowland Shrub** - Woody vegetation, less than 20 feet tall, with a tree cover of less than 10%, and occurring in wetland areas.

Broad-leaved Deciduous examples: Willow, Alder, Buckthorn
Broad-leaved Evergreen examples: Labrador-tea, Leather-leaf, Bog Rosemary
Coniferous examples: Stunted black spruce

c. **Forested Wetland** - Wetlands dominated by woody perennial plants, with a canopy cover greater than 10%, and trees reaching a mature height of at least 6 feet.

Coniferous examples: Black Spruce, Northern White Cedar, Tamarack
Broad-leaved Deciduous examples: Black Ash, Red Maple, Swamp White Oak
Mixed Broad-leaved Deciduous/Coniferous: Mixture of the species above. See Upland Mixed Broad-leaved Deciduous/Coniferous for group proportions.

Note: If an area meets the requirements of Forested Wetland, it should take precedence over any other "Forest" category.

VII. BARREN

Land of limited ability to support life and in which less than one-third (33-1/3%) of the area has vegetation or other cover. If vegetation is present, it is more widely spaced and scrubby than that in shrubland.

Note: If the area meets the requirements of both Agriculture and Barren, it should be placed in the Agriculture class. Also, if the area is wet and meets the requirements of Wetlands, it should be placed in the appropriate Wetland category.

a. **Sand**
b. **Bare Soil**
c. **Exposed Rock**
d. **Mixed** - an area that has less than two-thirds (66-2/3%) dominant cover of one of the above Barren classes.

VIII. SHRUBLAND

Upland Shrub - Vegetation with a persistent woody stem, generally with several basal shoots, low growth of less than 20 feet, and coverage of at least one-third (33-1/3%) of the land area. Less than 10% tree cover interspersed.

Examples: Scrub Oak, Buckthorn, Sumac

If the area is shrubland as a result of logging within the past 3 years, please indicate this in the comments section of the groundtruth sheet.

Note: See WETLAND (Lowland Shrub) for other shrub category

EXAMPLES

Below are some examples of how certain mixtures of forest are classified. An explanation is provided.

40% Maple, 10% Aspen, 5% Balsam Fir, 10% White Pine ......Broad-leaved Deciduous

This is called Broad-leaved Deciduous because there is one species that composes more than 33-1/3% of the canopy.

10% Aspen, 20% Maple, 10% Oak, 10% Balsam Fir, 15% Hemlock, 30% White Pine ..... Mixed Broad-leaved Deciduous/Coniferous

This is called Mixed Broad-leaved Deciduous/Coniferous because there are greater than 33-1/3% of each species group in the canopy.

35% Aspen, 20% Oak, 10% Balsam Fir, 20% White Pine, 5% Hemlock .....
Mixed Broad-leaved Deciduous/Coniferous

This is called Mixed Broad-leaved Deciduous/Coniferous because there are greater than 33-1/3% of each species group in the canopy, even though there is over 33-1/3% of Aspen.

20% Aspen, 80% Open Canopy with grasses in understory .....Broad-leaved Deciduous

This is called Broad-leaved Deciduous because only 10% canopy closure defines the forest class. A note on the groundtruth sheet should be made about the grass understory.

Appendix C. Methods for Reporting Accuracy Assessment Results

Note: The following document parallels and is based on sample data from the discussion of accuracy assessment in Lillesand and Kiefer (1994), pp. 615–618. For further information about these topics, please refer to that text.

The classification error matrix is a convenient and comprehensible method for displaying the results of the accuracy assessment process. Reference data are listed in the columns of the matrix and the classification data are listed in the rows. The major diagonal of the matrix represents the number of correctly classified samples; errors of omission are represented by the nondiagonal column elements, and errors of commission are represented by nondiagonal row elements. Table C.1 is an example of a classification error matrix, including six land cover categories.

Table C.1 Error matrix resulting from classification of random test pixels (based on Lillesand and Kiefer [1994], Table 7.4, p. 618).

	Reference Data						
	Water	**Sand**	**Forest**	**Urban**	**Corn**	**Hay**	**Row Total**
Water	226	0	0	12	0	1	239
Sand	0	216	0	92	1	0	309
Forest	3	0	360	228	3	5	599
Urban	2	108	2	397	8	4	521
Corn	1	4	48	132	190	78	453
Hay	1	0	19	84	36	219	359
Column Total	233	328	429	945	238	307	2840

Using the data from Table C.1, accuracy percentages can be calculated for the overall classification and for each category separately, as demonstrated in Table C.2. There are two distinct accuracy figures for the individual categories. The producer's accuracy is calculated by dividing the number of correctly classified samples by the column total for the category. The user's accuracy is calculated by dividing the number of correctly classified samples by the row total for the category.

Table C.2 Overall accuracy and producer's/user's accuracy by category.

Producer's Accuracy			**User's Accuracy**		
Water:	226/233 =	97.00%	Water:	226/239 =	94.56%
Sand:	216/328 =	65.85%	Sand:	216/309 =	69.90%
Forest:	360/429 =	83.92%	Forest:	360/599 =	60.10%
Urban:	397/945 =	42.01%	Urban:	397/521 =	76.20%
Corn:	190/238 =	79.83%	Corn:	190/453 =	41.94%
Hay:	219/307 =	71.34%	Hay:	219/359 =	61.00%

Overall accuracy = (226 + 216 + 360 + 397 + 190 + 219)/2,480 = 64.84%

Two-tailed 95% confidence intervals can be computed for the overall classification and for each category, as follows (Thomas and Allcock 1984; Jensen 1986; Snedecor and Cochran 1989):

$$CI = p \pm \left[1.96 \cdot \sqrt{pq \;/\; n} + (50 \;/\; n)\right]$$ [Equation 1]

where p = percent correct calculated above
q = 100 - p
n = number of samples

Table C.3 demonstrates the process of computing confidence intervals for overall accuracy and for category accuracy.

Table C.3 Computation of 95% confidence intervals (two-tailed) for overall accuracy and producer's/user's accuracy by category.

95% CI for overall accuracy:

$$64.84 \pm \left[1.96 \cdot \sqrt{64.84 \cdot 35.16 \;/\; 2480} + (50 \;/\; 2480)\right] = (62.94,\ 66.74)$$

95% CI for producer's accuracy by class:

Water: $97.00 \pm \left[1.96 \cdot \sqrt{97.00 \cdot 3.00 \;/\; 233} + (50 \;/\; 233)\right] = (94.60,\ 99.40)$

Sand: $65.85 \pm \left[1.96 \cdot \sqrt{65.85 \cdot 34.15 \;/\; 328} + (50 \;/\; 328)\right] = (60.57,\ 71.14)$

Forest: $83.92 \pm \left[1.96 \cdot \sqrt{83.92 \cdot 16.08 \;/\; 429} + (50 \;/\; 429)\right] = (80.32,\ 87.51)$

...

95% CI for user's accuracy by class:

Water: $94.56 \pm \left[1.96 \cdot \sqrt{94.56 \cdot 5.44 \;/\; 239} + (50 \;/\; 239)\right] = (91.48,\ 97.65)$

Sand: $69.90 \pm \left[1.96 \cdot \sqrt{69.90 \cdot 30.10 \;/\; 309} + (50 \;/\; 309)\right] = (64.63,\ 75.18)$

Forest: $60.10 \pm \left[1.96 \cdot \sqrt{60.10 \cdot 39.90 \;/\; 599} + (50 \;/\; 599)\right] = (56.10,\ 64.11)$

...

In addition to the figures provided in Tables C.2 and C.3, another measure of accuracy is widely used in accuracy assessment of land cover classifications. The Kappa, or KHAT, statistic describes the difference between the observed classification accuracy (represented by Table C.2) and the theoretical chance agreement that would result from a random classification (Congalton and Mead 1983; Rosenfield and Fitzpatrick-Lins 1986). For the overall classification, Kappa is computed as follows:

$$K = \frac{N \cdot \sum x_{ii} - \sum (x_{i+} x_{+i})}{N^2 - \sum (x_{i+} x_{+i})}$$ [Equation 2]

where N = total number of samples in all categories
$\Sigma (x_{ii})$ = number of correctly classified samples
$\Sigma (x_{i+} \cdot x_{+i})$ = sum of products of each category's row and column totals in the error matrix

For individual categories, this simplifies to the following:

$$K = \frac{N \bullet x_{ii} - x_{i+} x_{+i}}{N \bullet x_{i+} - x_{i+} x_{+i}}$$ [Equation 3]

where N = total number of samples in all categories
x_{ii} = number of correctly classified samples in the specified category
x_{i+} = row total in the error matrix for the specified category
x_{+I} = column total in the error matrix for the specified category.

The process of calculating Kappa statistics is demonstrated in Table C.4 below.

Table C.4 Kappa (KHAT) statistics for overall accuracy and category accuracy.

Kappa statistic for overall accuracy:

N = 2480 $\Sigma(x_{ii})$ = 226 + 216 + 360 + 397 + 190 + 219 = 1608
$\Sigma(x_{i+} \bullet x_{+I})$= (239•233) + (309•328) + (599•429) + (521•945) + (453•238) + (359•307) = 1,124,382
Kappa = {[(2480•1608) - 1,124,382] / [(2480•2480) - 1,124,382]} = 0.5697

Kappa statistic for category accuracy:

Water: Kappa = {[(2480•226) - (239•233)] / [(2480•239) - (239•233)]} = 0.9400
Sand: Kappa = {[(2480•216) - (309•328)] / [(2480•309) - (309•328)]} = 0.6532
Forest: Kappa = {[(2480•360) - (599•429)] / [(2480•599) - (599•429)]} = 0.5175
...

The variance of Kappa (Hudson and Ramm 1987) can be calculated as follows:

$$\sigma_K^2 = \frac{1}{N} \bullet \left[\frac{T(1-T)}{(1-U)^2} + \frac{2(1-T)(2TU-V)}{(1-U)^3} + \frac{(1-T)^2(W-4U)^2}{(1-U)^4} \right]$$ [Equation 4]

where $$T = \frac{\sum x_{ii}}{N}$$

$$U = \frac{\sum x_{i+} x_{+i}}{N^2}$$

$$V = \frac{\sum \left[x_{ii} \bullet (x_{i+} x_{+i}) \right]}{N^2}$$

$$W = \frac{\sum \sum \left[x_{ij} \bullet (x_{j+} x_{+i}) \right]}{N^3}$$

The process of calculating the variance of Kappa is demonstrated in Table C.5 below.

Table C.5 Kappa (KHAT) variance.

N = 2480

$\Sigma(x_{ii}) = 226 + 216 + 360 + 397 + 190 + 219 = 1608$

$\Sigma(x_{i+} * x_{+i}) = (239 \bullet 233) + (309 \bullet 328) + (599 \bullet 429) + (521 \bullet 945) + (453 \bullet 238) + (359 \bullet 307) = 1{,}124{,}382$

$\Sigma[x_{ii} * (x_{i+} + x_{+i})] = [226 \bullet (239+233)] + [216 \bullet (309+328)] + [360 \bullet (599+429)] +$

$[397 \bullet (521+945)] + [190 \bullet (453+238)] + [219 \bullet (359+307)] = 1{,}473{,}490$

$\Sigma[x_{ij} \bullet (x_{j+} + x_{+i})^2] = [226 \bullet (239+233)^2] + [0 \bullet (239+328)^2] + [3 \bullet 239+429)^2] + \ldots$

$\ldots + [78 \bullet (359+238)] + [219 \bullet (359+307)^2] = 2.279{,}167{,}222$

$T = (1608 / 2480) = 0.648387$

$U = [1{,}124{,}382 / (2480)^2] = 0.182814$

$V = [1{,}473{,}490 / (2480)^2] = 0.239576$

$W = [2{,}279{,}167{,}222 / (2480)^3] = 0.149424$

$\sigma^2(K) = (1/2480) \bullet [\ 0.341395 + -0.004595 + 0.004364\ ] = 0.0001376$

The Kappa statistic is often used to compare the results of multiple classifications (Congalton and Mead 1983; Congalton 1991). After calculating Kappa and its variance $\sigma^2(K)$ for each classification, a test statistic is computed as follows:

$$\frac{K_1 - K_2}{\sqrt{\sigma^2_{K1} + \sigma^2_{K2}}} \approx Z \qquad \text{[Equation 5]}$$

This test statistic follows a Gaussian (normal) distribution and can be used to determine whether differences between the two classifications are significant. Significance at 95% is obtained by comparing the Z-score to the equivalent value (1.96) from the normal tables. If the Z-score is greater than 1.96, the classification accuracy results are significantly different. The normal tables can also be used to test significance at other levels (e.g., 90%, 99%, or 99.9%) as desired.

This process is demonstrated in Table C.6 below.

Table C.6 Hypothesis test for comparing Kappa statistics.

Statistics from Classification 1:

$K_1 = 0.5697$ [from Table 4]
$\sigma^2(K_1) = 0.0001376$ [from Table 5]

Statistics from Classification 2:

$K_2 = 0.6024$
$\sigma^2(K_2) = 0.002539$

Statistics from Classification 3:

$K_3 = 0.6203$
$\sigma^2(K_3) = 0.0000794$

Threshold for significance at 95% = 1.96 [from normal tables]

$$Z_{2,1} = \frac{(0.6024 - 0.5697)}{\sqrt{0.002539 + 0.0001376}} = 0.6321$$ [not significant]

$$Z_{3,1} = \frac{(0.6203 - 0.5967)}{\sqrt{0.0000794 + 0.0001376}} = 3.4350$$ [significant]

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE June 1998	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE

Upper Midwest Gap analysis program image processing protocol

5. FUNDING NUMBERS

6. AUTHOR(S)

Thomas Lillesand, Jonathan Chipman, David Nagel, Heather Reese, Matthew Bobo, and Robert Goldmann

7. PERFORMING ORGANIZATION NAME AND ADDRESS

Environmental Remote Sensing Center, University of Wisconsin–Madison, 1225 West Dayton Street, Madison, Wisconsin 53706-1695, and Wisconsin Department of Natural Resources, PO Box 7921, Madison, Wisconsin 53707

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Geological Survey
Environmental Management Technical Center
575 Lester Avenue
Onalaska, Wisconsin 54650

10.SPONSORING/MONITORING AGENCY REPORT NUMBER

98-G001

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Release unlimited. Available from National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161 (1-800-553-6847 or 703-487-4650. Available to registered users from the Defense Technical Information Center, Attn: Help Desk, 8725 Kingman Road, Suite 0944, Fort Belvoir, VA 22060-6218 (1-800-225-3842 or 703-767-9050).

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This document presents a series of technical guidelines by which land cover information is being extracted from Landsat Thematic Mapper data as part of the Upper Midwest Gap Analysis Program (UMGAP). The UMGAP represents a regionally coordinated implementation of the national Gap Analysis Program in the states of Michigan, Minnesota, and Wisconsin; the program is led by the U.S. Geological Survey, Environmental Management Technical Center.

The protocol describes both the underlying philosophy and the operational details of the land cover classification activities being performed as part of UMGAP. Topics discussed include the hierarchical classification scheme, ground reference data acquisition, image stratification, and classification techniques. This discussion is primarily aimed at the image processing analysts involved in the UMGAP land cover mapping activities as well as others involved in similar projects. It is a "how-to" technical guide for a relatively narrow audience, namely those individuals responsible for the image processing aspects of UMGAP.

14. SUBJECT TERMS

Gap analysis, image processing protocol, Landsat, Michigan, Minnesota, Upper Midwest GAP, Wisconsin

15. NUMBER OF PAGES

25 pp. + Appendixes A–C

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	